OPAP IAP8 JUL 2 7 2011 PATENT & TRADEMARK OFFICE

944-003.182

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

| | | |
|---|---|---|
| Re Application of: Rämö et al. | : | Attorney Docket No.: 944-003.182 |
| Serial No.: 10/692,290 | : | Examiner: Michael N. Opsasnick |
| Filed: October 23, 2003 | : | Art Unit: 2626 |

**For: METHOD AND SYSTEM FOR SPEECH CODING**

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

## REPLY BRIEF OF APPELLANTS (37 CFR §41.41)

Sir:

This is a reply appeal in regard to the final rejection contained in a Final Office Action mailed on July 28, 2010, (the "Final Office Action"), and in reply to an Examiner's Answer (mailed May 26, 2011). This reply brief is filed within two months of the Examiner's Answer.

******If any fee and/or extension is required in addition to any enclosed herewith, please charge Account No. 23-0442.***

### CERTIFICATE OF MAILING/TRANSMISSION (37 C.F.R. § 1.8(a))

I hereby certify that this correspondence is, on the date shown below, being:

**MAILING**

☒ Deposited with the United States Postal Service with sufficient postage as first class Mail in an envelope addressed to the Commissioner for Patents, PO Box 1450, Alexandria, VA 22313-1450.

**FACSIMILE**

☐ Transmitted by facsimile to the U.S. Patent and Trademark Office.

Date: July 25, 2011

Marie E Forte
Signature
Marie E. Forte
*(type or print name of person certifying)*

## I. REAL PARTY IN INTEREST (37 CFR §41.37(c)(1)(i))

The real party in interest in this action is Nokia Corporation, Keilalahdentie 4, FIN-02150 Espoo, Finland, by virtue of the Assignment dated November 10 and 14, 2003. The Assignment was recorded in the U.S. Patent and Trademark Office on February 9, 2004, Reel 014970 and Frame 0234

## II. RELATED APPEALS AND INTERFERENCES (37 CFR §41.37(c)(1)(ii))

There are no related appeals or interferences.

## III. STATUS OF CLAIMS (37 CFR§41.37(c)(1)(iii))

The status of the claims is:

Claims pending: 1, 3-41 and 49-56

Claims canceled: 2, 42-48

Claims objected to: 15-18

Claims rejected: 1, 3-14, 19-41 and 49-56

Claims on appeal: 1, 3-14, 19-41 and 49-56

## IV. STATUS OF AMENDMENTS (37 CFR §41.37(c)(1)(iv))

No amendment of claims 1, 3-41 and 49-56 has been filed subsequent to final rejection.

## V. SUMMARY OF CLAIMED SUBJECT MATTER (37 CFR §41.37(c)(1)(v))

Appellant's invention is directed to a method and device related to the segmentation of an audio signal into a plurality of segments and the encoding of the segments with different encoding settings. The segmentation is chosen such that the intra-segment similarity of the speech parameters is high (page 11, lines 19-20). The segmentation can be made based on quantized or unquantized parameters (page 13, lines 4-5). After the segmentation, the segments can be classified into types so that each segment can be coded by a coding scheme based on the segment type (page 11, lines 18-25). In particular, the characteristics of the audio signal are indicated in the speech parameters extracted from a parameter extraction unit 12 (page 13, line 9-11), and the partitioning or segmentation is carried out by a compression module 20 based on the behavior of the parameters (page 13, lines 21-24).

The speech parameters are extracted at regular intervals including linear prediction coefficients, speech energy or gain, pitch and voicing information (page 11, lines 26-27). The pitch associated with the speech signal is shown in Figure 2b, the voicing information associated with the speech signal is shown in Figure 2c and the energy associated with the speech signal is shown in Figure 2d. The claimed invention uses those audio characteristics for partitioning the audio signal into a plurality of segments. For example, a segmentation algorithm can be implemented based on a number of audio characteristics (page 12, lines 1-24). An example of the audio signal segmentation, according to present invention, is shown in Figures 3a to 3d. Figure 3a shows an audio signal from frames 100 to frames 200. The energy associated with that audio signal is shown in Figure 3b and the voicing information associated with that audio signal is shown in Figure 3c. Based on the energy and the voicing information, the audio signal is segmented into 7 segments as shown in Figure 3d. Because the segments of the audio signal based on the audio characteristics will likely have different parameters associated with the audio characteristics, each segment can be efficiently coded using a coding scheme in order to meet the perceptual requirements, for example (page 12, lines 25-29). Thus, according to the claimed method, the partitioning of the audio signal is carried out based on the parameters indicative of the audio characteristics of the audio signal, and the segments are encoded with different encoding settings.

The invention of independent claim 1 is directed to a method for partitioning an audio signal into a plurality of segments based on parameters indicative of audio characteristics of the audio signal (page 12, lines 1-8; Figures 3a-3d; page 12, line 29 – page 13, line 7), the parameters are obtained from an audio signal for each of a plurality of consecutive time intervals (page 11, lines 26-27), and encoding the segments with different encoding settings (page 11, lines 24-25; page 12, lines 25-28).

The invention of independent claim 19 is directed to a decoder (item 40, Figure 4). The decoder comprises an input for receiving audio data and a module for generating a further audio signal (page 13, lines 15-20). The audio data is indicative of a plurality of segments obtained by partitioning the audio signal based on parameters indicative of audio characteristics of the audio signal, and extracted from each of a plurality of consecutive time intervals (Figures 3a-3d; page

12, line 29 – page 13, line 7; page 13, lines 9-11; page 11, lines 26-27). The audio data is also indicative of an adjusted representation of the parameters so that the further audio signal is generated based on the adjusted representation and the encoding settings (page 21, lines 21-26).

The invention of independent claim 22 is directed to an encoding device (item 20, Figure 4). The encoding device comprises an input for receiving audio data indicative of parameters (item 112, Figure 4), and an adjustment module for adjusting one or more of the parameters for providing an adjusted representation of the parameters, wherein the adjustment comprises partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals and encoding the segments based on one or more of a plurality of encoding settings (page 13, lines 8-12, lines 21-28; page 21, lines 21-26).

The invention of independent claim 27 is directed to an electronic device (item 40, Figure 4). The device comprises:

an input module for receiving audio data indicative of a plurality of segments of an audio signal (item 120, Figure 4), wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal (page 11, lines 26-27), and wherein the plurality of segments are obtained by partitioning the audio signal based on the parameters extracted for the consecutive time intervals (page 12, lines 1-8; Figures 3a-3d; page 12, line 29 – page 13, line 7; page 13, lines 9-11), and the audio data is indicative of the parameters in an adjusted representation (page 21, lines 21-32); and

a decoder, responsive to the audio data, for generating a synthesized audio signal based on the adjusted representation (page 13, lines 15-20).

The invention of independent claim 31 is directed to a communication network (Figure 11). The network comprises a plurality of base stations; and a plurality of mobile stations adapted for communicating with the base stations (page 23, line 26-31), wherein at least one of the mobile stations (item 50, Figure 4) comprises:

an input module for receiving audio data from at least one of the base stations, the audio data indicative of a plurality of segments of an input audio signal (item 120, Figure

4) , wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal (page 11, lines 26-27), and wherein the plurality of segments are obtained by partitioning the input audio signal based on the parameters extracted for the consecutive time intervals (page 12, lines 1-8; Figures 3a-3d; page 12, line 29 – page 13, line 7; page 13, lines 9-11), and encoded with a plurality of encoding settings based on the audio characteristics (page 12, lines 25-28), the audio data indicative of the parameters in an adjusted representation (page 21, lines 21-32).

a decoder, responsive to the audio data, for generating a synthesized audio signal based on the adjusted representation (page 13, lines 15-20).

In the invention of dependent claim 3, the audio characteristics include voicing characteristics in the segments of the audio signal (Figure 2c).

In the invention of dependent claim 4, the audio characteristics energy characteristics in the segments of the audio signal (Figure 2d).

In the invention of dependent claim 5, the audio characteristics include pitch characteristics in the segments of the audio signal (Figure 2b).

In the invention of dependent claim 6, the partitioning of the audio signal into segments is carried out concurrent to the encoding of the segments. As disclosed, a quantization mode is selected for each segmented parameter signal with *k* parameter values within the segment, but a reduced number *i* of parameters values are coded by the quantizer into the bitstream (page 15, lines 7-24).

In the invention of dependent claim 7, the partitioning is carried out before said encoding (page 7, lines 20-21).

In the invention of dependent claim 8, a plurality of voicing values are assigned to the audio characteristics of the audio signal in said segments, and the partitioning is carried out based on the assigned voicing values (page 11, lines 22-25).

In the invention of dependent claim 9, the plurality of values includes a value designated to a voiced speech signal and another value designated to an unvoiced signal (paragraph [0100]).

In the invention of dependent claim 10, the plurality of values further includes a value designated to a transitional stage between the voice and unvoiced signal (page 11, lines 28-31).

In the invention of dependent claim 11, the plurality of values further includes a value designated to an inactive period in the audio signal (page 11, lines 22-25).

In the invention of dependent claim 12, the encoding includes selecting a quantization mode for improving bit allocation and for reducing parameter update rate, and the partitioning is carried out based on the selected quantization mode (page 14, lines 27-31).

In the invention of dependent claim 13, the partitioning is carried out based on a selected target accuracy in reconstructing of the audio signal, wherein the target accuracy is selected based on a distortion criteria comparing upsampled quantized values and modified parameter signal (page 14, lines 30-35; page 16, lines 22-32).

In the invention of dependent claim 14, the partitioning also includes providing a linear pitch representation in at least some of the segments (page 21, lines 21-28).

In the invention of dependent claim 15, the audio signal is encoded into audio signal data, and the method further comprises: forming a parameter signal based on the audio signal data having a first number of signal data; downsampling the parameter signal to a second number of signal data for providing a further parameter signal, wherein the second number is smaller than the first number; and upsampling the further parameter signal to a third number of signal data in decoding, wherein the third number is greater than the second number (page 15, lines 13-33).

In the invention of dependent claim 16, the third number is equal to the first number (page 15, lines 25-27).

In the invention of dependent claim 17, the signal data comprises quantized parameters (page 13, lines 4-5).

In the invention of dependent claim 18, the signal data comprises unquantized parameters (page 13, lines 4-5).

In the invention of dependent claims 33, 37, 38, 39, 40, the encoding settings comprise bit allocation, quantization accuracy, quantization method and parameter update rate (Table II; page 19, lines 12-24; page 20, lines 16-20).

In the invention of dependent claim 34, the audio signal contains sinusoidal components and said parameters include frequency values, amplitude values and phase values indicative of the sinusoidal components (page 2, line 27 – page 3, line 11).

In the invention of dependent claim 35, the parameters include pitch, voicing, amplitude and energy of the audio signal (Figure 2).

In the invention of dependent claim 36, the parameters include pitch contour data containing a plurality of pitch values representative of an audio segment in time (Figure 10; page 21, lines 25-26).

In the invention of dependent claim 41, the audio signal comprises a plurality of frames and the audio signal in each frame has a waveform and wherein a further audio signal is produced in the decoding stage independently of the waveform (page 13, lines 8-20).

In the invention of dependent claims 49, 51, 53, the parameters are obtained from the audio signals in regular time intervals (page 11, lines 26-27).

In the invention of dependent claims 50, 52, 54, 55, 56 the partitioning is based on the similarity in the parameters among consecutive time intervals (page 11, lines 23-24).

The dependent claim 26 is directed to a computer readable storage medium embedded with a computer program having programming code for carrying out the method of claim 1 (Figure 4; page 13, lines 17-20).

In the invention of dependent claim 20, 28, the audio data is recorded on an electronic medium, and wherein input of the decoder is operatively connected to the electronic medium for receiving the audio data (Figure 4; page 13, lines 15-17).

In the invention of dependent claim 21, 29, the audio data is transmitted through a communication channel, and wherein the input of the decoder is operatively connected to the communication channel for receiving the audio data (page 13, lines 15-17).

In the invention of dependent claim 32, the parameters including pitch contour data containing a plurality of pitch values representative of an audio segment in time, and wherein the pitch contour data in the audio segment in time is approximated by a plurality of consecutive sub-segments in the audio segment for providing a plurality of end points, and wherein the end points include a first end point and a second end point for defining each of said sub-segments; and the decoder also includes a reconstruction module for reconstructing the audio segment based on the received audio data (Figure 10).

In the invention of dependent claim 23, the encoding device also comprises a quantization module, responsive to the adjusted representation, for coding the parameters in the adjusted representation (page 13, lines 15-20).

In the invention of dependent claim 24, the encoding device also comprises an output end, operatively connected to a storage medium, for providing data indicative of the coded

parameters in the adjusted representation to the storage medium for storage (Figure 4; page 13, lines 15-17).

In the invention of dependent claim 25, the encoding device also comprises an output end, operatively connected to a communication channel, for providing signals indicative of the coded parameters in the adjusted representation to the communication channel for transmission (Figure 4; page 13, lines 15-17).

In the invention of dependent claim 30, the electronic device comprises a mobile terminal (page 13, lines 15-17).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL (37 CFR §41.37(c)(1)(vi))

At section 2 of the Final Office Action, claims 1, 3-41, 49-56 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

In the Examiner's Answer, the 35 U.S.C. 112 rejection of claims 1, 3-41 and 49-56 has been withdrawn.

At section 3 of the Final Office Action, claims 1, 3-14, 19-21, 26-37, 39-44 and 46-48 are rejected under 102(b) as being anticipated by *Gersho et al.* (U.S. Patent No. 6,311,154, hereafter referred to as *Gersho*).

In particular, in rejecting independent claim 1, the Examiner states that *Gersho* discloses segmenting {partitioning or classifying} the audio input signal {speech} into a plurality of segments {frames} (partitioning samples of speech signal into frames, col.4,lines 25-27) based on the audio characteristics {classes} of the audio signal (classifying the speech signal in each frame into one of a plurality of classes, col.4, lines 25-27); and encoding the segments {frames} with different encoding settings {excitation} (encoding an excitation for the frame using one of the plurality of excitation coding ... selected according to the class of the frame, col.4, lines 30-33).

In rejecting independent claims 19 and 27 under 102(b), the Examiner states that *Gersho* discloses an input for receiving audio data indicative of parameters in the adjusted representation (Figure 3, input applied to filter 14), and a module for generating the audio signal based on the adjusted representation (Figure 3). The Examiner states that <u>it would have been inherent to one skilled in the art to use a decoder</u> to reverse the encoding data for further processing, such as modulating or storing the audio signal.

In rejecting independent claim 31, the Examiner states that *Gersho* discloses a cell phone system having both a base station and a mobile station (col.6, lines 33-36); a decoder (Figures 1, 4, 5, 9; col.8, lines 54-63); and an input for receiving audio data (Figures 1, 4, 5, col.3, lines 1-15).

At section 5, claims 15-18, 22-25, 38 and 45 are rejected under 102(e) as being anticipated by *Sinha et al.* (U.S. Patent No. 7,191,136 B2, hereafter referred to as *Sinha*).

In rejecting claims 15, 22, 23 and 45, the Examiner states that *Sinha* discloses a method for use in parametric audio coding to encode an audio signal by segmenting the audio signal, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, the one or more parameters relating to audio characteristics of the audio based on audio characteristics of the audio signal (by high-pass filtering the input audio signal as disclosed in col.4, lines 47-59) and then performing a non-linear parametric representation of the signal (col. 4, lines 53-59)).

In the Examiner's Answer, the 102(e) rejection of claims 15-18 has been withdrawn.

## VII. ARGUMENT (37 CFR §41.37(c)(1)(vii))

At issue here is whether the cited *Gersho* reference discloses partitioning an audio signal into a plurality of segments based on classes, and whether the terms "partitioning" and "classifying" are interchangeable.

In Section C below, applicant shows that *Gersho* does not disclose or suggest partitioning an audio signal into a plurality of segments based on classes. Applicant also contends that the terms "partitioning" and "classifying" are not interchangeable.

In Section L below, applicant shows that *Sinha* does not disclose or suggest partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals.

### A. The Claimed Invention

Claim 1 includes the limitations of

1) obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics of the audio signal,

2) partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals; and

3) encoding the segments with different encoding settings.

As pointed out in Section B below, the Examiner considers "parameters indicative of audio characteristics" as being equivalent to "classes", and "segments" as being equivalent to "frames".

Therefore, if claim 1 is anticipated by *Gersho*, *Gersho* must disclose or suggest partitioning the audio signal into a plurality of frames based on the classes obtained for the consecutive time intervals.

The cited *Gersho* reference does not disclose or suggest such limitation.

Each of the independent claims 19, 27 and 31 includes the limitation that the plurality of segments are obtained by partitioning the audio signal based on the parameters indicative of the audio characteristics of the audio signal (classes).

*Gersho* does not disclose or suggest that the plurality of segments are obtained by partitioning the audio signal based on classes.

B. 102 Rejection over *Gersho*

In rejecting claim 1, the Examiner states that *Gersho* teaches:

segmenting {**partitioning** or **classifying**} the audio signal into a plurality of segments {**frames**} (partitioning samples of a speech signal into frames, col.4, lines 25-27) and

obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics {**classes**} of the audio signal (classifying the speech signal in each frame into one of the plurality of classes, col.4, lines 25-27).

It is respectfully submitted that the terms "partitioning" and "classifying" are not interchangeable. The term "partitioning", when applied to a speech frame, means dividing the speech frame into smaller units such as sub-frames. The term "classifying", when applied to a speech frame, means designating the speech frame as "a voiced frame" or "an unvoiced frame", for example. Therefore, the 102 rejection of claim 1 can be split into two versions:

Version A:

classifying the audio signal into a plurality of frames; and

obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of classes of the audio signal.

Version B:

partitioning the audio signal into a plurality of frames (partitioning samples of a speech signal into frames, col.4, lines 25-27) and

obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of classes of the audio signal (classifying the speech signal in each frame into one of the plurality of classes, col.4, lines 25-27).

Version A is irrelevant to the claimed invention. Version B does not read on the limitation of claim 1 because the partitioning of the audio signal into a plurality of frames is not based on classes.

C. The Cited *Gersho* Reference

*Gersho* does not disclose or suggest partitioning the audio signal into a plurality of segments based on the classes obtained for the consecutive time intervals.

In col. 4, lines 23-27, *Gersho* discloses a method for coding a speech signal includes the steps of

a) partitioning samples of a speech signal into **frames**;

b) deriving a residual signal for each frame; and

c) classifying the speech signal in each **frame** into one of a plurality of classes.

A person skilled in the art would understand that *Gersho* performs "partitioning samples of a speech signal into frames" **before** "classifying the speech signal in each frame into one of a plurality of classes". As *Gersho* performs "classifying the speech signal in each frame" **after** the step of partitioning, the information on classes is not available at the time of partitioning. Therefore, *Gersho* does not disclose the limitation of partitioning the audio signal into a plurality of segments **based on** the classes obtained for the consecutive time intervals.

D. The Examiner's Answer Regarding *Gersho*

In the Examiner's Answer, the Examiner introduces **new** argument using steps d) and e) and Figure 2 to show that *Gersho* teaches the claimed limitations. In particular, the Examiner states:

*Of essence, is the performance of steps d) and e), where the subframe size is **resized** based upon the previous processing (classifying) of the audio signal. See Figure 2, where the "basic sub-frame", sized as n2-n1, is adjusted to either a minimum of n2-n1 – (d2+d2) or a maximum n2-n1+(d2_d1); nonetheless, the subframe **size is readjusted after** the measurement of parameters of the input signal. Or to match claim words, "time intervals" of claim 1 map to Gersho's frame, and the "partitioning of segments" map to Gersho's resizing of the subframe.* *See* page 10, third paragraph of the Examiner's Answer, with emphasis by the Examiner.

Applicant respectfully disagrees.

First, in the Examiner's Answer, the Examiner considers "resizing" as being equivalent to "partitioning" and "sub-frames" as being equivalent to "segments". This assertion is **inconsistent** with the 102 rejection of claim 1 (*see* Section B above).

Second, *Gersho* does not disclose or suggest resizing the sub-frames based on classes.

On col.4, lines 28-32, *Gersho* discloses:

d) identifying the location of a least one window in the frame by examining the residual signal for the frame; and

e) encoding an excitation for the frame using one of a plurality of excitation coding techniques selected according to the class of the frame.

In step d, *Gersho* discloses identifying the **window** location in the frame by examining the **residual signal** for the frame. The **residual signal** is obtained in step b, which has nothing to do with the class information as obtained in a later step c.

In step e, *Gersho* discloses using class information for encoding an excitation for the frame, but not for the partitioning the audio signal into frames.

Furthermore, according to *Gersho*, **windows** are selected time intervals within the sub-frame such that the excitation signal within a sub-frame is constrained to be zero outside the windows. *See* col. 3, lines 26-29. Figure 2 only shows how a search sub-frame is associated with a basic sub-frame, for the purpose of performing LP analysis on the input speech and for the purpose of packaging the data to be transmitted into a fixed number of bits for each fixed frame interval. *See* col. 7, lines 18-27. As shown in Figure 2, a basic sub-frame extends from n1 to n2. One may associate a basic sub-frame with a search frame which extends from n1+d1 to n2+d2. The magnitudes of d1 and d2 are defined so as to be always less than half the window size, and their values are chosen so that each search sub-frame will contain an integer number of windows. *See* col.8, lines 12-23. Based on the length of the sub-frames, windows are established so that all of the non-zero excitation amplitudes located within the windows. *See* col. 7, lines 51-56. In particular, for the AbS search process, the sub-frame size is adaptively modified to assure that an integer number of windows will be present in the excitation segment to be coded. *See* col. 8, lines 8-11.

Thus, *Gersho* only discloses locating one or more windows within a frame such that all the non-zero excitation amplitude is located within the windows. The locating is based on examining the residual signal. This has nothing to do with partitioning the audio signal into a plurality of segments based on classes.

In sum, *Gersho* does not disclose or suggest 1) obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics of the audio signal, and 2) partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals.

E. 102 Rejection of Claims 1, 19, 27 and 31

As pointed out in Sections B to D above, *Gersho* does not disclose or suggest

obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics of the audio signal, and

partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals.

For the above reasons, *Gersho* fails to anticipate independent claim 1.

*Gersho* does not disclose or suggest that the plurality of segments are obtained by partitioning the audio signal based on the parameters indicative of the audio characteristics of the audio signal.

For the above reasons, *Gersho* fails to anticipate independent claims 19, 27 and 31.

F. Dependent Claims 4, 6, 12, 13, 33, 34, 37, 39, 40, 41, 50, 52, 54, 55 and 56

It is respectfully submitted that claims 4, 6, 12, 13, 33, 34, 37, 39, 40, 41, 50, 52, 54, 55 and 56 are dependent from claims 1, 19, 27 and 31 and include further limitations. For reasons regarding claims 1, 19, 27 and 31 above, *Gersho* fails to anticipate claims 4, 6, 12, 13, 33, 34, 37, 39, 40, 41, 50, 52, 54, 55 and 56. In addition, the Examiner does not clearly point out where *Gersho* discloses the further limitations in claims 4, 6, 12, 13, 33, 34, 37, 39, 40, 41, 50, 52, 54, 55 and 56.

F.1 102 Rejection of Claim 4

In rejecting claim 4, the Examiner states that *Gersho* discloses that the characteristics include energy characteristics (energy in the **residual signal**) in said segments {windows} of the audio signal. Col. 4, lines 65-67.

It is respectfully submitted that claim 4 is dependent from claim 1. Claim 4 includes the limitations of

obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics (classes) of the audio signal, and

partitioning the audio signal into a plurality of segments (frames) based on the parameters obtained for the consecutive time intervals, wherein the characteristics include energy characteristics in said segments of the audio signal.

It is respectfully submitted that the residual signal is obtained in step b) which is derived in each frame **after** the speech samples are partitioned in step a) into frames, and before the speech signal in each frame is classified in step c) into one of a plurality of classes. *See* Section C above.

*Gersho* does not disclose or suggest that the residual signal is part of the parameters on which the partitioning of audio signals is based. For this reason alone, *Gersho* fails to anticipate claim 4.

F.2 102 Rejection of Claim 6

In rejecting claim 6, the Examiner states that *Gersho* discloses segmenting {partitioning} is carried out concurrently {classifying and encoding} to said encoding {coding} (partitioning samples of speech, classifying speech signals into classes, coding a speech signal, col.4, line 24-25. The Examiner states that the classifying and encoding process may be done concurrently.

First, claim 4 includes the limitation that **segmenting** is carried out concurrently to said **encoding**. As pointed out in Section B above, the terms "classifying" and "partitioning" are not interchangeable. Thus, whether *Gersho* discloses that classifying and encoding process may be done concurrently is irrelevant to the claimed invention.

Second, in *Gersho,* the segmenting is carried out in step a) and the encoding is carried out in step e), whereas classifying is carried out in step c). *Gersho* does not disclose or suggest that the partitioning and encoding process may be done concurrently.

For this reason alone, *Gersho* fails to anticipate claim 6.

F.3 102 Rejection of Claim 12

In rejecting claim 12, the Examiner states that *Gersho* discloses that said encoding comprises selecting a quantization mode for improving bit allocation and for reducing parameter update rate, wherein the partitioning is carried out based on the selected quantization mode (col.3, lines 45-49; Figure 5 and col.11, lines 4-16, col.4, lines 36-37, col.15, lines 35-36 and col.9, lines 63-65).

In col. 3, lines 45-49, *Gersho* discloses:

> *In accordance with a further aspect of this invention a highly efficient encoding of the excitation frame is achieved by directing processing to the windows themselves, and allocating all or nearly all of the available bits to code the regions inside the windows.*

The above passage has nothing to do with "selected quantization mode".

Figure 5 only shows that the encoder has two stages, an adaptive codebook first stage and a ternary pulse coder second stage. In the first stage, a segment of the past of the excitation signal is selected as the first approximation to the excitation signal in the subframe. The second stage 26 is based on a ternary pulse coding method where the coder identifies three non-zero pulses, one selected from the sample positions 0, 3, 6, 9, 12, 15, 18, 21; the second pulse position is selected from 1, 4, 7, 10, 13, 16, 19, 22, and the third pulse from 2, 5, 8, 11, 14, 17, 20, 23. Thus three bits are needed to specify each of the three pulse positions, and one bit is needed for the polarity of each pulse. *See* col.10, lines 45-58.

Figure 5 has nothing to do with "selected quantization mode".

In col. 11, lines 4-16, *Gersho* discloses:

*The location of each window in each basic frame of voiced speech is determined by the energy contour peaks and is transmitted to the decoder. An improved performance can be obtained if the location is found by performing the AbS search process for each candidate location, but this technique results in higher complexity. A fixed window size of 24 samples is used with only one window per search subframe. Three bits are used to specify the starting point of each window using a quantized time grid, i.e., the start of a window is allowed to occur at multiples of 8 samples. In effect, the window location is "quantized", thereby reducing the time resolution with a corresponding reduction in the bit rate.*

In the above passage, *Gersho* only discloses how the window location is quantized. The passage does not suggest "partitioning of the frames is carried out based on the selected quantization mode".

In col.4, lines 35-39, *Gersho* discloses:

*In one embodiment the classes include voiced frames, unvoiced frames, and transition frames, while in another embodiment the classes include strongly periodic frames, weakly periodic frames, erratic frames, and unvoiced frames.*

In col.15, lines 35-36, *Gersho* discloses:

*If the Rate(m)=1, then the current frame is declared as a silent frame. If not, (i.e. if Rate(m)=3 or 4), the current frame is declared as active speech.*

In col. 9, lines 63-65, *Gersho* discloses:

*Referring to FIG. 4, a frame classifier 22 sends two bits per basic frame to the speech decoder 10 (see FIG. 14) in the receiver to identify the class (00, 01, 10, 11).*

The above three passages have nothing to do with "selected quantization mode".

*Gersho* does not disclose or suggest that the encoding comprises selecting a quantization mode for improving bit allocation and for reducing parameter update rate, wherein the partitioning is carried out based on the selected quantization mode. For the above reason alone, *Gersho* fails to anticipate claim 12.

F.4 102 Rejection of Claim 13

In rejecting claim 13, the Examiner states that *Gersho* discloses that said partitioning is carried out based on a selected target accuracy in reconstructing of the audio signal, wherein the target accuracy is selected based on a distortion criteria comparing upsampled quantized values (transmitted samples) and modified parameter signal (col.9, lines 63-65 and col.3, lines 45-49).

In col. 9, lines 63-65, *Gersho* discloses:

*Referring to FIG. 4, a frame classifier 22 sends two bits per basic frame to the speech decoder 10 (see FIG. 14) in the receiver to identify the class (00, 01, 10, 11).*

In the above passage, *Gersho* only discloses using two bits to identify the class of the basic frame. This has nothing to do with the target accuracy on which partitioning is based.

In col.3, lines 45-49, *Gersho* discloses:

*In accordance with a further aspect of this invention a highly efficient encoding of the excitation frame is achieved by directing processing to the windows themselves, and allocating all or nearly all of the available bits to code the regions inside the windows.*

In the above passage, *Gersho* only discloses a method for increasing the encoding efficient of the frame. This has nothing to do with the target accuracy on which partitioning is based.

For the above reason alone, *Gersho* fails to anticipate claim 13.

F.5 102 Rejection of Claims 33, 37, 39 and 40

In rejecting claims 33, 37, 39 and 40 the Examiner states that *Gersho* teaches that the encoding settings comprise bit allocation (col.3, lines 45-49), quantization accuracy (Figure 5

and col.11, lines 4-16), quantization method (col.11, lines 4-16) and parameter update rate (col.3, lines 31-44 and 56-60).

Figure 5 only shows that the encoder has two stages, an adaptive codebook first stage and a ternary pulse coder second stage. In the first stage, a segment of the past of the excitation signal is selected as the first approximation to the excitation signal in the subframe. The second stage 26 is based on a ternary pulse coding method where the coder identifies three non-zero pulses, one selected from the sample positions 0, 3, 6, 9, 12, 15, 18, 21; the second pulse position is selected from 1, 4, 7, 10, 13, 16, 19, 22, and the third pulse from 2, 5, 8, 11, 14, 17, 20, 23. Thus three bits are needed to specify each of the three pulse positions, and one bit is needed for the polarity of each pulse. *See* col.10, lines 45-58.

Figure 5 has nothing to do with quantization or quantization accuracy.

In col. 11, lines 4-16, *Gersho* discloses:

> *The location of each window in each basic frame of voiced speech is determined by the energy contour peaks and is transmitted to the decoder. An improved performance can be obtained if the location is found by performing the AbS search process for each candidate location, but this technique results in higher complexity. A fixed window size of 24 samples is used with only one window per search subframe. Three bits are used to specify the starting point of each window using a quantized time grid, i.e., the start of a window is allowed to occur at multiples of 8 samples. In effect, the window location is "quantized", thereby reducing the time resolution with a corresponding reduction in the bit rate.*

In the above passage, *Gersho* only discloses how the window location is quantized. The passage does not suggest the quantization method in the encoding setting for encoding the frame.

In col.3, lines 31-44, *Gersho* discloses:

> *In accordance with a further aspect of this invention there is disclosed a technique for determining the location and size of the windows, and identifying those critical segments of the excitation signal which are particularly important to represent with a suitable*

*selection of pulse amplitudes. The subframe and frame sizes are allowed to vary (in a controlled manner) to suit the local characteristics of the speech signal. This provides for an efficient coding of the windows without having a window cross a boundary between two adjacent subframes. In general, the size of the windows and their locations are adapted according to the local characteristics of the input or target speech signal. As employed herein, locating a window refers to positioning a window around energy peaks associated with the residual signal, depending on the short-term energy profile.*

In the above passage, *Gersho* only discloses how to locate the windows within a frame in order to increase the coding efficiency. This passage has nothing to do with parameter update rate – the rate in which the parameters indicative of the audio characteristics (classes) are updated.

In col.3, lines 56-60, *Gersho* discloses:

*A toll quality speech coding technique in accordance with this invention is a time-domain scheme which exploits novel ways to represent and encode speech signals at different data rates, depending on the nature and the amount of information contained in short-time segments of the speech signal.*

In the above passage, *Gersho* only discloses that the data rates for encoding speech signals are depending on the information contained in the time segments of the speech signal. This passage has nothing to do with parameter update rate – the rate in which the parameters indicative of the audio characteristics (classes) are updated.

*Gersho* does not disclose or suggest all the limitations in claims 33, 37, 39 and 40. For the above reasons alone, *Gersho* fails to anticipate claims 33, 37, 39 and 40.

F. 6 102 Rejection of Claim 34

In rejecting claim 34, the Examiner states that *Gersho* teaches that the audio signal contains sinusoidal components (col.3, lines 25-29, analysis windows made equal become sine) and said parameters include frequency values (Figure 1, element 68), amplitude values (col.3,

lines 51-55) and phase values indicative of the sinusoidal components (Figure 1, element 76 and col.3, lines 25-29)..

First, the expression "**analysis windows made equal become sine**" is incomprehensible. What is sine? Does the expression mean that the width of the windows is made equal to the wavelength of the speech signal? *Gersho* does not suggest such a window.

Second, the width of the window, according to *Gersho*, is described as follows:

In col. 3, lines 25-29, *Gersho* discloses:

> *In accordance with one aspect of this invention the excitation signal within a subframe is constrained to be zero outside of selected intervals within the subframe. These intervals are referred to herein as windows.*

The selection of window width does not suggest that the input signal contains sinusoidal components.

In Figure 1, element 68 is a frequency synthesizer which provides the required frequencies to the receiver (col.6, lines 41-44). Having a frequency synthesizer does not suggest that the parameters indicative of the audio characteristics (classes) include frequency values.

In col.3, lines 51-55, *Gersho* discloses:

> *Further in accordance with the teachings of this invention, a reduced complexity method for coding the signal inside a window is based on the use of ternary valued amplitudes, 0, -1, and +1. The reduced complexity coding method is also based on exploiting a correlation between successive windows in periodic speech segments.*

In the above passage, *Gersho* only discloses a method of using three ternary amplitude values to represent the signal inside a window in order to reduce the encoding complexity. However, the ternary amplitude values are not the same as the amplitude values (of the sinusoidal components of the input signal) included in the parameters indicative of the audio characteristics (classes).

In Figure 1, element 76 is an IQ demodulator which drives in-phase (I) and quadrature (Q) signals from the received signal. Having an IQ demodulator does not suggest that the phase values of the sinusoidal components are included in the parameters indicative of the audio characteristics (classes).

In col. 3, lines 25-29, *Gersho* discloses:

*In accordance with one aspect of this invention the excitation signal within a subframe is constrained to be zero outside of selected intervals within the subframe. These intervals are referred to herein as windows.*

In the above passage, *Gersho* only discloses how the windows are defined. This passage does not suggest that the phase values of the sinusoidal components are included in the parameters indicative of the audio characteristics (classes).

*Gersho* does not disclose or suggest all the claim limitations of claim 34. For this reason alone, *Gersho* fails to anticipate claim 34.

F.7 102 Rejection of Claim 41

In rejecting claim 41, the Examiner states that *Gersho* discloses that the audio signal comprises a plurality of frames and the audio signal in each frame has a waveform and wherein a further audio signal is produced in the decoding stage independently of the waveform (col. 14, lines 8-14; col.13, lines 63-67 and col. 14, 1-7).

In col.13, line 63 to col.4, line 14, *Gersho* discloses:

*To achieve an efficient excitation representation, and in accordance with an aspect of this invention described previously, the fixed codebook contribution in a voiced frame is constrained to be zero outside of selected intervals (windows) within that frame. The separation between two successive windows in voiced frames is constrained to be equal to one pitch period. The locations and sizes of the windows are chosen so that they jointly represent the most critical segments of the ideal fixed codebook contribution. This technique, which focuses the attention of the encoder on the perceptually important*

*segments of the speech signal, ensures efficient encoding.*

*A voiced frame is typically divided into three subframes. In an alternative embodiment, two subframes per frame has been found to be a viable implementation. The frame and subframe length may vary (in a controlled manner). The procedure for determining these lengths ensures that a window never straddles two adjacent subframes.*

Nothing in the above passages suggests that the audio signal produced in the decoding stage is independent of the waveform in each of the frames.

For this reason alone, *Gersho* fails to anticipate claim 41.

F. 8 102 Rejection of claims 50, 52, 54, 55 and 56

In rejecting claims 50, 52, 54, 55 and 56, the Examiner states that *Gersho* teaches regular and consecutive time intervals (Figure 2 and Figure 6).

It is respectfully submitted that claims 50, 52, 54, 55, 56 include the limitations that the partitioning is based on the similarity in the parameters among consecutive time intervals.

In *Gersho*, Figure 2 only shows how a search sub-frame is defined. Figure 6 shows the three pulse positions 1, 2, and 3 based on a ternary pulse coding method (col. 10, lines 49-63).

*Gersho* does not disclose or suggest that the partitioning of the audio signal is based on the similarity in the parameters (indicative of the audio characteristics or classes) among consecutive time intervals.

For the above reason alone, *Gersho* fails to anticipate claims 50, 52, 54, 55 and 56.

G. Dependent Claims 3, 5, 7-11, 14, 20, 21, 26, 28-30, 32, 35, 36, 49, 51 and 53

Claims 3, 5, 7-11, 14, 20, 21, 26, 28-30, 32, 35, 36, 49, 51 and 53 are dependent from claims 1, 19, 22, 27 and 31 and include further limitations. For reasons regarding claims 1, 19, 27 and 31 above, *Gersho* also fails to anticipate claims 3, 5, 7-11, 14, 20, 21, 26, 28-30, 32, 35, 36, 49, 51 and 53.

H. 102 Rejection over *Sinha*

At section 5 of the Final Office Action, claims 15-18, 22-25 and 38 are rejected under 35 U.S.C. 102(e) as being anticipated by *Sinha*. In the Examiner's Answer, the rejection of 15-18 has been withdrawn.

I. Independent Claim 22

Claim 22 includes the limitations of:

an input for receiving audio data indicative of parameters obtained from an audio signal in a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal; and

an adjustment module for adjusting one or more of the parameters for providing an adjusted representation of the parameters, wherein said adjusting comprises partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals and encoding the segments based on one or more of a plurality of encoding settings.

J. 102 Rejection of Claim 22

In the Final Office Action, in the rejection of claims 22, 23 and 45, the Examiner states that *Sinha* teaches a method for use in a parameter audio coding to encode an audio signal by

segmenting the audio signal, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters relating to audio characteristics of the audio signal (col.4, lines 47-51, by high-pass filtering the input audio signal); performing a non-linear parameter representation of the signal; col. 4, lines 53-59 – wherein the data amount per processing depends upon the frequency characteristics of the audio signal, and the characteristics analyzed can be peak analysis, lattice quantization, or frequency range selection – col.3, lines 1-6); and encoding the segments with different encoding settings (by choosing compression settings on-the-fly – col.6, lines 43-47)

The Examiner only states that *Sinha* discloses segmenting the audio signal, for each of a plurality of consecutive time intervals, one or more parameters.

The Examiner fails to cite *Sinha* for disclosing partitioning the audio signal into a plurality of segments based on the parameters as claimed.

K. The Examiner's Answer Regarding *Sinha*

On page 17 of the Examiner's Answer, the Examiner states that applicant's claim scope does not pertain to "partitioning into segments" because the compression block 20 in Figure 4 of the patent application is disqualified.

Applicant respectfully submitted that the **112 rejection** of claims 1, 3-41 **has been withdrawn**. *See* page 4 of the Examiner's Answer. This indicates that claims 22-25 and 38 are properly disclosed.

Furthermore, that the compression block 20 is used for segmenting is described on page 13, lines 21-24 of the specification.

In the Examiner's Answer, the Examiner states that *Sinha* teaches a method for use in a parametric audio coding to encode an audio signal by segmenting the audio signal into a plurality of segments based on audio characteristics of the audio signal (by high-pass filtering the input audio signal; by performing a non-linear parametric representation of the signal; wherein the data amount per processing depending upon the frequency characteristics of the audio signal, and the characteristics analyzed can be peak analysis, lattice quantization or frequency range selection, col. 4, lines 47-51, col.4, lines 53-59 col.3, lines 1-6).

In col.2, line 66, *Sinha* discloses:

> *The present invention also allows for compression mechanisms to be determined "on-the-fly" and transmitted via the header at playback time. The type of features which may be adaptively chosen include techniques such as lattice quantization of scale factors, multidimensional coding of the peaks, and selection of a frequency range most amenable towards efficient high frequency coding.*

In the above passage, *Sinha* only discloses that the encoder adaptively chooses one or more of the lattice quantization of scale factors, multidimensional coding of peaks, or frequency range for efficiency high frequency coding. *See* col. 9, lines 48-52. In particular, lattice quantization of scale factor is a process to decode the Huffman Scale Factor using the lattice codebooks or non-lattice codebooks (see col.7, line 14-18); multidimensional coding of peaks is a process to decode the spectrum peaks using the multidimensional peaks (see col.7, lines 18-23). The Examiner fails to specifically point out which of these high frequency coding

techniques is equivalent to segmenting the audio signal into a plurality of segments based on audio characteristics of the audio signal.

In col.4, lines 43-59, *Sinha* discloses:

*The above-described enhancement of the present invention is outlined in FIG. 4. In this coding scheme the compressed information consists of coded low frequency components (from the low pass filter 402 with a cut-off frequency of $f_l$) as well as a parametric representation for the high frequency components (from the high pass filter 404 with a cut-off frequency of $f_h$): based on a non-linear model 406. The parametric representation requires significantly fewer bits than conventional coding of the higher frequency components. These parameters for the non-linear high frequency model representation are updated every audio frame (an audio frame in PAC typically consists of 1024 PCM samples). Next, the non-linear model parameters 408 estimated for the non-linear model 406 (using a method described below) are then combined with standard PAC coded output (via a PAC encoder 410) to form the encoded output of the audio signal.*

In the above passage, *Sinha* only discloses a coding scheme which compresses information consisting of coded low frequency components as well as parametric representations for the high frequency components from the high pass filter (Abstract; column 4, lines 44-49). In particular, *Sinha* allows the input signal to pass through both a high pass filter and a low-pass filter so that the audio components in the high-frequency range and the audio components in the low-frequency range are encoded using different models. The parametric representation for the high frequency components is estimated based on a non-linear model.

As known to a person skilled in the art, filtering out the high or low frequency components in an audio signal is not equivalent to segmenting the audio signal into a plurality of segments. The Examiner also fails to point out whether the coding of the low frequency components or the coding of parametric representation for the high frequency components are equivalent to segmenting the audio signal into a plurality of segments based on audio characteristics of the audio signal.

Furthermore, *Sinha* does not disclose or suggest partitioning the audio signal into a plurality of segments <u>based on the parameters</u> obtained for the consecutive time intervals.

L. The Cited *Sinha* Reference

According to *Sinha*, segmentation of the speech signal into frames takes place before the high-pass and low-pass filtering and before the non-linear parameters 408 are obtained. *Sinha* does not disclose or suggest partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals.

*Sinha* discloses a method for improving an audio compression scheme, such as perceptual audio coding (PAC). In a conventional PAC scheme, as shown in Figure 1, the input signal is segmented into frames to be stored in a frame buffer 104. The frames are then processed through a long-term predictor 106 and a short-term predictor 108 for linear predictive analysis (col. 2, lines 13-16). Each of the audio frames in PAC consists of 1024 pulse code modulated (PCM) samples (col.3, lines 52-56). According to *Sinha*, as the input speech signal is segmented into frames of 1024 PCM samples, the speech signal is simultaneously provided to a low-pass filter 402 for obtaining compressed information consisting of coded low frequency components, and to a high-pass filter 404 for obtaining a parametric representation of the high frequency components based on a non-linear model 406. The parametric representation is updated every audio frame in order to estimate the non-linear parameters 408 (col.4, lines 43-59).

According to *Sinha*, segmentation of the speech signal into frames takes place before the high-pass and low-pass filtering and before the non-linear parameters 408 are obtained.

Thus, *Sinha* does not disclose or suggest partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals.

M. *Sinha* Fails to Anticipate Claim 22

As pointed out in Sections K and L above, *Sinha* does not disclose or suggest partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals. For the above reasons, *Sinha* fails to anticipate claim 22

N. 102 Rejection of Claim 23-25 and 38

It is respectfully submitted that claims 23-25 and 38, they are dependent from claim 22 and include further limitations. For reasons regarding claim 22 above, *Sinha* also fails to anticipate claims 23-25 and 38.

## VIII CLAIMS APPENDIX (37 CFR §41.37(c)(1)(viii))

1. A method, comprising:

obtaining, for each of a plurality of consecutive time intervals, one or more parameters from an audio signal, said one or more parameters indicative of audio characteristics of the audio signal,

partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive time intervals; and

encoding the segments with different encoding settings.

2. (canceled)

3. A method according to claim 1, wherein the characteristics include voicing characteristics in said segments of the audio signal.

4. A method according to claim 1, wherein the characteristics include energy characteristics in said segments of the audio signal.

5. A method according to claim 1, wherein the characteristics include pitch characteristics in said segments of the audio signal.

6. A method according to claim 1, wherein said partitioning is carried out concurrent to said encoding.

7. A method according to claim 1, wherein said partitioning is carried out before said encoding.

8. A method according to claim 1, wherein a plurality of voicing values are assigned to the audio characteristics of the audio signal in said segments, and wherein said partitioning is carried out based on the assigned voicing values.

9. A method according to claim 8, wherein the plurality of values includes a value designated to a voiced speech signal and another value designated to an unvoiced signal.

10. A method according to claim 8, wherein the plurality of values further includes a value designated to a transitional stage between the voice and unvoiced signal.

11. A method according to claim 8, wherein the plurality of values further includes a value designated to an inactive period in the audio signal.

12. A method according to claim 1, wherein said encoding comprises selecting a quantization mode for improving bit allocation and for reducing parameter update rate, wherein the partitioning is carried out based on the selected quantization mode.

13. A method according to claim 1, wherein said partitioning is carried out based on a selected target accuracy in reconstructing of the audio signal, wherein the target accuracy is selected based on a distortion criteria comparing upsampled quantized values and modified parameter signal.

14. A method according to claim 5, wherein said partitioning comprises providing a linear pitch representation in at least some of said segments.

15. A method according to claim 1, wherein the audio signal is encoded into audio signal data, said method further comprising:

forming a parameter signal based on the audio signal data having a first number of signal data;

downsampling the parameter signal to a second number of signal data for providing a further parameter signal, wherein the second number is smaller than the first number; and

upsampling the further parameter signal to a third number of signal data in decoding, wherein the third number is greater than the second number.

16. A method according to claim 15, wherein the third number is equal to the first number.

17. A method according to claim 15, wherein the signal data comprises quantized parameters.

18. A method according to claim 15, wherein the signal data comprises unquantized parameters.

19. A decoder, comprising:

an input for receiving audio data indicative of a plurality of segments of an audio signal, wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal, and wherein the plurality of segments are obtained by partitioning the audio signal based on the parameters extracted for the consecutive time intervals, and the audio data is indicative of the parameters in an adjusted representation; and

a module, responsive to the audio data, for generating a further audio signal based on the adjusted representation and the encoding settings.

20. A decoder according to claim 19, wherein the audio data is recorded on an electronic medium, and wherein input of the decoder is operatively connected to the electronic medium for receiving the audio data.

21. A decoder according to claim 19, wherein the audio data is transmitted through a communication channel, and wherein the input of the decoder is operatively connected to the communication channel for receiving the audio data.

22. An encoding device comprising:

an input for receiving audio data indicative of parameters obtained from an audio signal in a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal; and

an adjustment module for adjusting one or more of the parameters for providing an adjusted representation of the parameters, wherein said adjusting comprises partitioning the audio signal into a plurality of segments based on the parameters obtained for the consecutive

time intervals and encoding the segments based on one or more of a plurality of encoding settings.

23. An encoding device according to claim 22, further comprising a quantization module, responsive to the adjusted representation, for coding the parameters in the adjusted representation.

24. An encoding device according to claim 22, further comprising an output end, operatively connected to a storage medium, for providing data indicative of the coded parameters in the adjusted representation to the storage medium for storage.

25. An encoding device according to claim 22, further comprising an output end, operatively connected to a communication channel, for providing signals indicative of the coded parameters in the adjusted representation to the communication channel for transmission.

26. A computer readable storage medium embedded with a computer program comprising programming code for carrying out the method of claim 1.

27. An electronic device comprising:

an input module for receiving audio data indicative of a plurality of segments of an audio signal, wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal, and wherein the plurality of segments are obtained by partitioning the audio signal based on the parameters extracted for the consecutive time intervals, and the audio data is indicative of the parameters in an adjusted representation; and

a decoder, responsive to the audio data, for generating a synthesized audio signal based on the adjusted representation.

28. An electronic device according to claim 27, wherein the audio data is recorded in an electronic medium, and wherein the input is operatively connected to the electronic medium for receiving the audio data.

29. An electronic device according to claim 27, wherein the audio data is conveyed through a communication channel, and wherein the input is operatively connected to the communication channel for receiving the audio data.

30. An electronic device according to claim 27, comprises a mobile terminal.

31. A communication network, comprising:

a plurality of base stations; and

a plurality of mobile stations adapted for communicating with the base stations, wherein at least one of the mobile stations comprises:

an input module for receiving audio data from at least one of the base stations, the audio data indicative of a plurality of segments of an input audio signal, wherein one or more parameters are extracted from the audio signal for each of a plurality of consecutive time intervals, the parameters indicative of audio characteristics of the audio signal, and wherein the plurality of segments are obtained by partitioning the input audio signal based on the parameters extracted for the consecutive time intervals and encoded with a plurality of encoding settings based on the audio characteristics, the audio data indicative of the parameters in an adjusted representation; and

a decoder, responsive to the audio data, for generating a synthesized audio signal based on the adjusted representation.

32. A decoder according to claim 19, the parameters including pitch contour data containing a plurality of pitch values representative of an audio segment in time, and wherein the pitch contour data in the audio segment in time is approximated by a plurality of consecutive sub-segments in the audio segment for providing a plurality of end points, and wherein the end points include a first end point and a second end point for defining each of said sub-segments; and

a reconstruction module for reconstructing the audio segment based on the received audio data.

33. A method according to claim 1, wherein the encoding settings comprise bit allocation, quantization accuracy, quantization method and parameter update rate.

34. A method according to claim 1, wherein the audio signal contains sinusoidal components and said parameters include frequency values, amplitude values and phase values indicative of the sinusoidal components.

35. A method according to claim 1, wherein the parameters include pitch, voicing, amplitude and energy of the audio signal.

36. A method according to claim 1, wherein the parameters include pitch contour data containing a plurality of pitch values representative of an audio segment in time.

37. A decoder according to claim 19, wherein the encoding settings include bit allocation, quantization accuracy, quantization method and parameter update rate.

38. An encoding device according to claim 22, wherein the encoding settings include bit allocation, quantization accuracy, quantization method and parameter update rate.

39. A computer readable storage medium according to claim 26, wherein the encoding settings include bit allocation, quantization accuracy, quantization method and parameter update rate.

40. A communication network according to claim 31, wherein the encoding settings include bit allocation, quantization accuracy, quantization method and parameter update rate.

41. A method according to claim 1, wherein the audio signal comprises a plurality of frames and the audio signal in each frame has a waveform and wherein a further audio signal is produced in the decoding stage independently of the waveform.

Claims 42-48. (canceled)

49. A method according to claim 1, wherein the parameters are obtained from the audio signals in regular time intervals.

50. A method according to claim 1, wherein said partitioning is based on the similarity in the parameters among consecutive time intervals.

51. A decoder according to claim 19, wherein the parameters are extracted from the audio signals in regular time intervals.

52. A decoder according to claim 19, wherein the plurality of segments are obtained based on similarity in the parameters among consecutive time intervals.

53. An encoding device according to claim 22, wherein the parameters are obtained from the audio signals in regular time intervals.

54. An encoding device according to claim 22, wherein said partitioning is based on similarity in the parameters among consecutive time intervals.

55. An electronic device according to claim 27, wherein the plurality of segments are obtained based on similarity in the parameters among consecutive time intervals.

56. A communication network according to claim 31, wherein said partitioning is based on similarity in the parameters among consecutive time intervals.

## IX. EVIDENCE APPENDIX (37 CFR §41.37(c)(1)(ix))

There are no evidences submitted pursuant to 37 CFR §1.130, 1,131 or 1,132.

## X. RELATED PROCEEDING APPENDIX (37 CFR §41.37(c)(1)(x))

There are no prior decisions rendered by a court or the Board in any proceeding identified pursuant to 37 CFR §41.37(c)(1)(ii).

## CONCLUSION

It is respectfully submitted that the present invention as claimed is readily distinguishable over the cited *Gersho* reference. Appellants' invention is not disclosed in the applied prior art and there is no fair basis for alleging that appellants' invention is obvious in regard to such art.

In view of the above, it is respectfully submitted that the rejection of claims 1, 3-16, 19-41 and 49-56 is in error and must be reversed.

Respectfully submitted,

Date: July 25, 2011

Ken Lao

Kenneth Q. Lao
Attorney for the Appellant
Registration No. 40,061

WARE, FRESSOLA, VAN DER SLUYS
& ADOLPHSON LLP
Bradford Green, Building Five
755 Main Street, P.O. Box 224
Monroe, CT 06468
Telephone: (203) 261-1234
Facsimile: (203) 261-5676
USPTO Customer No. 004955